Desk

Vol. 42 DECEMBER, 1951 No. 8

The Journal of Educational Psychology

ERSITY
OF MICHIGAN
JAN 15 1 ?
PERIODICAL
READING ROOM

CONTENTS

Title Page and Index, Volume 42 (1951) . . . . . i–viii

Patterns in the Mass Communications Tastes of the Young Audience 449
PAUL I. LYNESS

An Analysis of a Procedure Used in the Teaching and Rating of Oral Expression 468
FRANCIS J. DI VESTA AND JAMES H. L. ROACH

The Effect of Attention-gaining Devices on Film-mediated Learning 479
D. MORGAN NEU

A Comparison of Reading Ability and Readability. 491
JOHN U. MICHAELIS AND FRED T. TYLER

Notes on a Suggested Index of Item Validity: The U-L Index. . . 499
A. PEMBERTON JOHNSON

Published monthly except June to September
$6.00 a year in the United States, Pan American Countries and the Philippines; Canada, $6.20; other countries, $6.40
Single issues $1.10

WARWICK & YORK, INC.
BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as Second Class Matter at York, Pa.

THE JOURNAL OF Educational Psychology

Established 1910

EDITORIAL BOARD

STEPHEN M. COREY
Teachers College, Columbia University
—*Learning*—

J. B. STROUD
State University of Iowa
—*Individual Differences*—

JACK W. DUNLAP
10 E. 49th St., New York 17, N.Y.
—*Technical, Statistics*—

PERCIVAL M. SYMONDS
Teachers College, Columbia University
—*Mental Hygiene*—

KARL J. HOLZINGER
University of Chicago
—*Factor Analysis*—

MILES A. TINKER
University of Minnesota
—*Reading*—

HAROLD E. JONES
University of California
—*Social Behavior, Child Psychology*—

ALEXANDER G. WESMAN
The Psychological Corporation
—*Tests*—

H. H. REMMERS
Purdue University
—*Attitudes, Teacher Evaluation*—

PAUL A. WITTY
Northwestern University
—*Children's Interests*—

H. E. BUCHHOLZ
Managing Editor
10 E. Centre St., Baltimore 2, Md.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. The JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

Books and other materials for review and correspondence regarding editorial matters should be addressed to The Journal of Educational Psychology, Warwick and York, Inc., Publishers, 10 E. Centre St., Baltimore 2, Md.

Manuscripts should be typed and double-spaced throughout, including quotations, footnotes, and references. In order to attain

(Continued on Inside Back Cover)

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 42 December, 1951 Number 8

PATTERNS IN THE MASS COMMUNICATIONS TASTES OF THE YOUNG AUDIENCE*

PAUL I. LYNESS†

Development Division, The Curtis Publishing Company, Philadelphia

This article summarizes a recent study of the mass communications habits and tastes of young people in school. The major goal of the study was to discover what 'patterns of interest' in the content of the mass media as a whole characterized five different age groups of boys and girls. By 'patterns of interest' we mean here the tendency for certain subject matter to cluster in the content preferences of any given age-sex group, and the tendency for these preferences to carry over from medium to medium.[1]

* The study of which this article is a summary was the author's doctoral dissertation in mass communications, written under the direction of Professor Leslie G. Moeller, Director of the School of Journalism, State University of Iowa.

† Formerly head of the Research Division, School of Journalism, State University of Iowa.

[1] Alice Sterner (*Radio, Motion Picture and Reading Interests; A Study of High-School Pupils*, Contributions to Education, No. 932, Teachers College, Columbia University, 1947) has provided us with one approach to this problem of communication interest patterns. She strove to answer these questions: "What really impels youth to spend so much time (in mass communications activities?) . . . Is it the medium itself or a major adolescent interest in adventure, humor or love, which attracts the pupil?" Some other researchers who have studied interests in all or nearly all the media simultaneously are: Wilbur Schramm, "Reading and Listening Patterns of American University Students," *Journ. Quarterly*, March, 1945; Walter J. Scott, *Reading, Film and Radio Tastes of High-School Boys and Girls*, Wellington, New Zealand, 1947; Sidney L. Pressey and Francis P. Robinson, *Psychology and the New Education*, New York, 1944, pp. 126–139; Frederick J. Meine, "Radio and Press among Young People," in Lazarsfeld and Stanton, *Radio Research–1941*, New York, 1941, pp. 189–223; Paul A. Witty and Anne Coomer, "Activities and Preferences of a Secondary-

It was felt that if rather detailed patterns of mass communications interest could be established and described meaningfully, editors, station managers, and producers would find them useful in 'beaming' their output to specific audiences of young people. Such information would also enable teachers better to understand their pupils and to utilize the mass media more effectively in their teaching programs. Social researchers, too, preoccupied with the prediction and control of behavior, might find the material valuable in testing the interaction between youthful mass communications interests and personality, adjustment and behavior.

With this as the problem, then, a rather comprehensive study was designed and carried out in the public schools of Des Moines, Iowa, in March, 1950.[2]

METHOD

Samples in five grades—third, fifth, seventh, ninth and eleventh—were selected by essentially random methods, although an initial stratification by sex and school attended was made. Students above the third grade filled out questionnaires in the classroom. Third-grade students were interviewed individually. The questionnaires were administered by teachers in the Des Moines public school system and graduate students at the State University of Iowa, all of whom were carefully briefed on the purposes of the survey and the procedures to be followed. The questions covered the amount and kind of reading, radio listening and movie-going; the preferred types of mass communications content (by medium); the names of preferred newspaper columnists, radio programs, movies, magazines, books and comic books; and preferred media.

A series of six pretests involving pupils in the third, fourth, fifth, seventh, eleventh and twelfth grades permitted a practical evaluation of the study design, the questionnaires and the methods employed.

School Group," *Journal of Educational Psychology*, February, 1943, pp. 65–76.

[2] The study was in large part financed by the *Des Moines Register and Tribune Company*. The writer wishes to thank Henry Kroeger, Director of Research and *The Iowa Poll, Des Moines Register and Tribune*, for his extensive personal assistance in carrying out the project.

CHARACTERISTICS OF THE SAMPLES

The total sample of 1,418 school children was made up of six hundred ninety-one boys and seven hundred twenty-seven girls. Over three hundred students in each of grades V, VII, IX and XI took part in the study. In addition, one hundred fifteen third-grade students, fifty-nine boys and fifty-six girls, were interviewed.

The breakdown of the total, by occupation of father, showed the following distribution: professional, eleven per cent; managers and proprietors, fifteen per cent; clerical and sales, twenty-two per cent; service, six per cent; skilled labor, twenty-two per cent; semi-skilled labor, eighteen per cent; unskilled labor, three per cent; other, three per cent.

By education of father, the percentage breakdown for fifth-grade students was: grade school, 5.6; high school, 26.6; college, 35.6; don't know, 32.2. For seventh-grade students, the breakdown in percentage (following the same order as above) was 4.6, 47.4, 26.1 and 21.9 respectively; and for the ninth- and eleventh-grade respondents, approximately 13.5, 44.3, 28.5, and 13.6. Information on the education of parents was not obtained from third-grade students.

Other background data obtained in the survey included ownership of automobiles, telephones, radios and TV sets; church affiliation and attendance; years of residence in Des Moines; preferred presidential candidate in 1948; possession of a library card; and recreational activities at home and away from home.

REPRESENTATIVES OF THE SAMPLES

The samples were substantially representative of their respective grade populations, in terms of sex and school attended, by virtue of the initial stratification. Figures on parental education and occupation for all public school children were not available. However, an estimate of 'representativeness' in terms of these factors was made as follows: A statistical comparison was made of educational and occupational data gathered by *The Iowa Poll*[3] on city adults with children in school and city adults without

[3] *The Iowa Poll* is a weekly, state-wide survey of what Iowans think on leading topics of the day. It is sponsored by *The Des Moines Register and Tribune.*

children in school (state-wide surveys). The differences between these groups were insignificant at the five per cent level of confidence.

Then, the educational and occupational data in the present survey were compared with three recent *Iowa Polls* of adults in the city of Des Moines. This test revealed no significant differences in the occupational categories (at the five per cent level of confidence), except in 'Managers and Proprietors,' where the difference was significant at the two per cent level. The present samples contained more people from middle and lower status homes than *The Iowa Poll* samples. On the other hand, the test showed that the present samples were somewhat overloaded with children whose parents had attended high school and college.

THE SURVEY CITY

Des Moines, at the time of the survey, had a population of approximately 176,954 (1950 census). The morning *Register* and the evening *Tribune*, both published by the *Des Moines Register and Tribune Company*, were the city's daily newspapers.

All four major networks had outlets in Des Moines, and, in addition, two independent stations were located there. In 1949, there were twenty-four moving picture theatres in the city. Des Moines citizens had access to over half a million library books in the ten public libraries, the Drake University library and the State Traveling Library.

THE QUESTIONS RELATING TO PREFERRED TYPES OF CONTENT

In the questionnaire given to fifth-, seventh-, ninth- and eleventh-grade pupils, seven questions dealt with preferred types of subject matter in the various communications media. One question was devoted to each of the following: newspapers, radio, moving pictures, types of magazines, magazine stories and features, books, and types of news.

A check list of content categories accompanied each question. The pupils were asked to select and rank their four favorite types of content from each list. Thus, the question relating to motion pictures was phrased as follows: "11. Here are several different types of movies. Which four do you usually like best? Write '1' before the type of movie you like best. Write '2'

before your second choice. Write '3' and '4' before your third and fourth choices."

The check list of motion picture categories appeared below the question (See Table III).

The radio and magazine check lists contained several names of programs and magazines exemplifying the categories. The illustrative programs had high Hooperatings in Des Moines at the time of the study and represented the various networks and stations heard in Des Moines. For example, "Western shows like The Lone Ranger, Roy Rogers, Tom Mix, Hopalong Cassidy, etc.," and "Radio columnists like Jimmy Fidler, Louella Parsons, Walter Winchell, Sheila Graham, etc." The magazines named were high circulation magazines in each classification. For example, "Sports magazines like *Sports Afield*, *Outdoor Life*, *Sport*, etc.," and "Girls' magazines like *Senior Prom*, *The American Girl*, etc."

In determining the preferences for each sex-grade group, the first, second, third and fourth choices for each medium were weighted and combined. (First choices were weighted 4; second choices, 3; third choices, 2; and fourth choices, 1.) The weighted scores for each medium were then converted to percentages. This made easier the comparison of preferences in the different media.

FINDINGS

Tables I–VII contain the content ratings for each sex-grade group, and for all media, plus types of news.

It should be borne in mind that these ratings are based upon the weighted first, second, third and fourth choices, and not upon a single favorite in each medium. The four-choice approach was used in order to get some measure of intensity or depth of interest.

A tabulation of first choices alone disclosed high group interest in a few categories. For example, the number one newspaper favorites of about three-fourths of all respondents in all grades were either comics or sports. With only a quarter of the first choices remaining, to be distributed among fifteen to twenty categories, the ranking of the secondary interests became somewhat difficult and imprecise. On the other hand, the tabulation of four weighted choices revealed strong secondary interests in

TABLE I.—FAVORITE PARTS OF THE NEWSPAPER
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
2.2	2.1	1.4	2.0	Advertisements	.8	1.9	1.0	1.2
.1	.4	.2		Book Reviews	1.0	.2	.5	.1
.6	.7	.3	.3	Business pages	.1	.4		.1
1.4	1.1	2.0	2.9	Columns on public affairs	.9	.9	2.1	2.8
.5	.2	.8	.9	Columns on love & personal problems	3.1	4.8	5.7	4.8
33.5	30.4	24.7	24.8	Comics	34.4	34.7	26.8	22.9
4.8	3.2	1.0	1.1	Continued stories	5.3	4.0	1.1	.3
1.4	.5	2.0	2.2	Editorials	.6	1.2	1.0	2.4
1.2	1.1	1.4	1.1	Gossip Col's	1.0	2.0	3.4	3.0
6.2	6.7	7.8	5.8	Headlines	8.0	5.9	7.1	6.0
.2	.3	.5	.4	Letters to editor	.4	.7	.8	1.5
11.1	10.0	7.8	4.3	Movies or radio news	12.2	13.2	8.8	5.8
4.7	9.5	12.6	15.0	News (general)	7.6	6.0	10.6	10.8
8.6	7.2	7.8	8.0	Pictures	12.1	9.8	7.5	5.7
		.1		Society; Women's features	1.6	3.0	5.6	12.2
22.7	26.4	28.1	30.5	Sports	9.8	10.7	17.0	20.3
.7	.2	1.4	.7	Other	1.0	.7	1.0	.1
.1		.1		None	.1			
1456	1616	1547	1410	Weighted total	1557	1945	1458	1576

such newspaper categories as movie and radio news, women's features, and general news, depending upon grade and sex. Comparisons of first choices with the combined four choices did show considerable correspondence in the direction and magnitude of changes from grade to grade.

One important qualification to the findings contained in Tables I–VII should be considered. The rating of each content category for each sex-grade group is a relative rating; that is, it is

relative to all other categories for a given medium. Thus, the fact that interest in newspaper comics fell off somewhat with age, relative to other newspaper categories, does not mean that the older children's absolute interest in comics was less. It merely

TABLE II.—FAVORITE TYPES OF RADIO PROGRAMS
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
14.2	11.2	4.3	1.7	Adventure serials	9.4	6.5	2.9	1.1
13.7	15.1	19.1	18.1	Comedy shows	13.5	15.4	16.7	13.3
.1	.4	1.1	2.3	Classical music	1.2	1.1	2.1	5.0
.3	.5	.1	.1	Love serials	3.6	1.8	1.8	.8
12.1	18.0	17.3	17.4	Mystery plays	12.0	18.2	16.7	11.7
1.0	2.8	2.6	3.4	News	1.3	.6	2.0	2.6
8.5	4.0	5.3	3.8	Oldtime and west. music	7.7	5.8	2.9	2.6
6.0	7.5	9.6	10.0	Other plays (radio drama)	12.0	18.7	20.3	21.5
.5	2.5	6.8	15.8	Popular music	3.2	7.8	15.1	23.3
8.4	7.6	8.3	4.2	Quiz shows	12.3	10.4	7.9	5.5
.6	.7	1.1	1.5	Radio columnists	.5	.6	1.4	1.8
.4	.6	1.1	1.5	Religious programs	1.2	1.9	2.5	2.1
8.9	12.4	13.7	15.5	Sports	2.7	3.0	4.1	6.5
.2	.1	.3	.2	Talks; Forums		.1		.3
24.8	16.2	9.0	3.6	Western shows	19.2	7.4	3.1	1.1
.3	.4	.3	.9	Other	.2	.7	.5	.8
1573	1632	1574	1379	Weighted total	1581	1930	1470	1587

means that the relative rank of comics among other newspaper categories was lower among the older boys and girls.

Several questions in the study were designed to get from the boys and girls information on what they actually read, heard on the radio and saw in the movies. The responses to these questions indicated a sizeable increase in the variety of reading, listening and viewing from the third through the eleventh grades. For instance, analysis of 'regularly read' newspaper content

showed an increased or stable proportion of responses, from grade to grade, in every category except 'book reviews' and 'continued stories.' Coming back to the example above, between 95 and 100 per cent of the boys and girls in all grades said they read the comics regularly.

TABLE III.—FAVORITE TYPES OF MOTION PICTURES
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
11.1	14.8	18.4	19.6	Adventure	9.2	9.6	9.8	9.4
14.4	8.0	5.4	3.7	Cartoons	16.8	6.9	3.0	1.9
14.7	15.5	13.5	15.2	Comedy features	14.5	13.8	12.7	9.7
2.1	4.3	4.1	6.2	Historical movies	.7	.7	3.0	5.0
.7	4.9	5.1	2.7	Horror movies	1.0	3.7	3.2	2.0
1.3	1.0	2.2	2.3	Love and romance	8.0	12.9	12.8	14.8
2.2	1.7	6.9	11.7	Musicals	11.7	18.1	22.2	28.4
9.1	10.3	12.4	11.5	Mystery	9.5	10.3	13.5	8.2
.4	1.0	2.2	1.2	Religious movies	1.9	3.2	4.9	4.3
.3	.3	.8	1.5	Social problems	.8	.7	1.2	3.5
14.8	17.5	15.5	13.7	War	1.8	2.6	3.6	4.5
28.2	18.1	12.2	8.9	Westerns	18.6	11.8	5.9	2.5
.7	1.8	1.2	1.4	World of make believe	4.4	4.5	3.0	3.4
	.8	.1	.4	Other drama	1.1	1.2	1.2	2.4
1339	1567	1559	1340	Weighted total	1365	1877	1457	1519

In radio program types listened to regularly, there was an increased or stable proportion of responses, from grade to grade, in all categories, except quiz shows, westerns, adventure serials, old-time and western music, and love serials.

Keeping in mind these reservations, one may make three general statements with respect to the data contained in Tables I–VII. 1) Both grade and sex influenced the degree of preference for certain types of content. 2) Preferences for certain types of content tended to remain at about the same level from age group to age group. 3) The tables disclose at least three

hypothetical 'patterns of interest' or generalized interests, which may be described as (a) violence and adventure; (b) love, private life, and glamor; and (c) 'educational' content (content which may be described as primarily informative and serious).

TABLE IV.—FAVORITE TYPES OF MAGAZINES
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
3.7	3.8	4.6	1.3	Adventure	2.0	1.6	.6	.1
12.4	15.8	12.8	6.2	Boys'	.4	.1	.3	.1
9.7	8.8	6.1	3.7	Detective story	7.6	5.3	3.0	.4
4.0	2.7	2.4	1.4	Farm	2.2	.3	.3	.5
2.6	3.9	11.7	14.4	General appeal	4.4	7.3	11.9	13.5
.5	.2	.4		Girls'	11.4	16.6	15.7	15.1
1.5	.9	.9	1.8	Home service	6.2	4.6	3.3	6.3
	.4	.1	.1	Literary	.1		.2	
4.1	4.9	5.6	9.0	Men's	.8	.4	.1	.7
3.4	3.6	2.8	1.6	Movie	12.6	16.4	14.3	9.0
1.1	1.4	1.3	2.8	News	.7	.5	.2	.9
5.5	7.8	9.5	16.5	Picture	9.4	7.9	9.2	11.2
8.1	9.4	12.4	14.1	Science	1.9	.6	.5	.2
19.1	20.6	21.7	23.3	Sports	6.8	3.2	3.0	2.3
.8	1.3	1.5	1.2	Religious	1 2	1.8	2.0	1.2
1.0	1.1	1.2	1.0	Romance	8.5	12.3	11.5	7.2
22.0	13.0	4.8	1.6	Western	13.5	5.8	2.3	.3
.5	.1			Women's	8.2	8.4	9.1	10.3
	.3	.2		Young women's	2.1	6.9	12.5	20.7
1312	1562	1560	1364	Weighted total	1354	1887	1441	1580

I. INFLUENCE OF GRADE AND SEX ON PREFERENCES FOR CERTAIN TYPES OF CONTENT

Those effects of grade and sex which were especially marked will be discussed under two major headings:

A) Categories which rose in relative popularity with the grade level of the respondents:

Both sexes.—The preference level of both boys and girls rose steadily, from the fifth through the eleventh grades, in the case of sports and general news in the newspaper; radio drama, popular music and sports on the radio; movie musicals and historical movies; general magazines; sports stories and features in magazines; and international and foreign news in types of news.

Popular music, radio drama, and movie musicals were consistently more popular with the girls than the boys. Newspaper sports and general news, radio and magazine sports, and international news were consistently more popular with boys than girls.

Boys.—The boys' preference level, but not the girls', increased with respect to the following: movie adventure; science and invention in magazines, books, and types of news; men's magazines; picture magazines; adventure books and books on sports. All these categories were more popular with boys than girls, except for picture magazines and adventure books which were about equally popular with the sexes.

Girls.—The girls' preference level, but not the boys', increased with respect to the following: women's features and society, and gossip columns in the newspaper; classical music on the radio; love and romance, religion, and war in the movies; love and romance, and women's subjects in magazines; magazines for women and young women; historical fiction, and love and romance in books. All these categories, except war and historical fiction, were more popular with girls than boys.

B) Categories which fell off in relative popularity as the grade level of the respondents rose:

Both sexes.—The preference level of both boys and girls dropped, from the fifth through the eleventh grades, in the case of comics, movie and radio news in the newspaper; westerns in radio, movies, magazines and books; adventure serials, old-time and western music, and quiz shows on the radio; movie cartoons; detective story magazines; and auto wrecks and Hollywood in types of news. Westerns and adventure serials were consistently more popular with boys than girls. Movie and radio news was consistently more popular with the girls.

Boys.—The boys' preference level, but not the girls', fell off with respect to boys' magazines and movie magazines; and crime

TABLE V.—FAVORITE TYPES OF MAGAZINE STORIES AND FEATURES
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
15.2	13.3	13.6	17.2	Adventure	12.4	9.4	8.6	6.1
2.2	1.1	.5	1.4	Art and music	6.1	2.5	2.3	3.0
.4	.3	.2	.6	Business and industry	.1	.1		.1
6.1	7.4	7.8	5.0	Cartoons	8.2	8.6	6.7	5.4
8.4	8.1	9.0	6.7	Crime; Police; gangsters	4.9	4.6	4.7	1.2
5.4	3.3	4.5	5.0	History and hist'l fiction	1.2	1.2	2.1	3.9
2.6	2.2	1.2	1.8	Hollywood and movies	8.9	10.7	11.3	7.6
.4	.1	.1	.4	Home service	.4	1.1	.1	1.4
10.1	11.8	12.5	10.6	Humor and comedy	13.9	15.9	16.6	17.3
	.3	.5	.2	Letters to editor	.4	.6	1.6	.7
.8	.9	2.0	.9	Love and romance	6.6	13.5	13.9	18.8
10.1	9.8	7.8	8.0	Mystery; detective	8.2	9.6	9.0	4.4
.1	1.0	1.5	1.4	Public affairs		.1	.2	1.3
.3	1.2	1.6	.9	Religion and churches	2.7	2.0	2.9	2.1
2.8	4.3	8.4	9.2	Science and invention	1.3	1.4	1.7	1.3
.6	.5	1.3	.4	Social problems		.1	.7	2.4
10.5	16.1	17.8	22.9	Sports	3.7	4.1	5.6	8.9
.2	1.7	1.4	2.4	Travel	1.0	.7	1.2	2.7
22.4	13.9	7.0	4.1	Westerns	14.7	6.1	2.3	1.2
	.1			Women's subjects	1.6	5.0	7.0	9.0
1.4	2.0	.4	.2	World of make believe	3.7	2.4	1.5	1.0
	.6	.9	.7	Other		.3		.2
1242	1556	1519	1334	Weighted total	1359	1790	1454	1557

in types of news. Boys' magazines and crime news were more popular with boys than girls. Movie magazines were more popular with girls than boys.

TABLE VI.—FAVORITE TYPES OF BOOKS
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
12.8	15.1	19.2	18.3	Adventure	14.6	16.2	14.9	15.8
2.7	1.1	.6	1.5	Art or music	8.5	3.0	2.1	3.2
.8	.9	2.2	1.2	Biography	2.7	1.4	2.1	3.9
6.1	4.1	5.0	5.5	History and hist'l fiction	1.0	2.0	3.6	9.2
.2	.4	.5	.1	Home service	.6	1.0	1.7	1.4
13.4	11.8	12.0	13.0	Humor and comedy	14.8	20.4	21.0	20.6
.5	.7	1.5	1.2	Love and romance	8.5	13.6	15.6	18.1
9.8	10.2	8.8	7.5	Mystery; detective	10.1	13.7	13.1	5.9
.2	.3	.8	.9	Public affairs	.1	.1	.1	
	.9	1.0	2.2	Religious books	2.1	2.6	3.7	3.2
3.9	5.6	8.3	7.1	Science and invention	3.0	2.7	1.6	1.6
.2	.2	.1		Social problems	.1	.2	.8	1.7
13.1	16.8	19.2	22.0	Sports	7.1	5.6	8.5	6.9
.6	1.8	1.3	1.9	Travel	1.9	1.2	1.9	1.9
11.1	12.2	8.0	11.4	War	1.2	1.2	.9	1.5
22.1	14.7	8.3	5.0	Westerns	17.3	8.1	4.0	2.0
2.1	1.6	1.3	.4	World of make believe	6.1	4.8	2.6	2.3
.4	1.6	1.9	.8	Other	.3	2.2	1.8	.8
1319	1595	1561	1391	Weighted total	1429	1882	1489	1566

Girls.—The girls' preference level, but not the boys', fell off with respect to magazine adventure; the world of make believe in magazines and books; and sports magazines. Magazine adventure and sports magazines were more popular with boys than girls. The world of make believe was more popular with girls than boys.

TABLE VII.—FAVORITE TYPES OF NEWS
(In Per cents of Weighted 1st, 2nd, 3rd and 4th Choices)

Boys					Girls			
5th	7th	9th	11th		5th	7th	9th	11th
16.3	13.6	8.9	6.9	Auto wrecks; accidents	9.3	13.4	9.2	6.4
5.4	1.9	1.2	1.6	Art, lit., and music	12.0	5.4	3.8	5.9
.1	.3	1.1	1.0	Business and industry	.2	.1	.7	.6
18.9	17.3	16.3	13.6	Crime; police; gangsters	10.8	11.1	12.6	8.9
8.1	9.8	8.6	8.4	Des Moines news	13.0	10.5	12.9	14.5
9.6	7.9	8.0	9.1	Disasters	6.9	9.7	10.6	9.2
3.4	7.3	10.4	7.7	Internat'l and foreign news	3.9	3.9	7.0	7.5
.5	.7	2.2	1.5	Labor Prob.	.7	.3	.5	1.1
1.2	.2	.7	.1	Love and marriage problems	5.9	4.3	3.4	4.2
1.6	2.0	3.3	4.2	Military affairs	.2	.6	.3	.3
6.4	5.8	3.0	2.5	Movies and Hollywood	16.8	19.0	14.5	9.6
1.1	.6	2.5	3.3	Politics and gov't	.8	.9	1.2	2.7
1.1	2.2	1.6	1.3	Religion and churches	2.8	3.7	5.1	3.0
4.6	5.9	8.0	10.2	Science and invention	3.1	3.2	1.8	3.7
.8	.7	1.1	.3	Social problems	.9	1.1	1.6	2.9
20.6	23.6	22.8	28.3	Sports	12.7	12.2	14.5	19.1
.3	.2	.3		Other		.6	.3	.4
1368	1609	1577	1388	Weighted total	1438	1909	1480	1580

II. CONTENT PREFERENCES WHICH TENDED TO REMAIN AT ABOUT THE SAME LEVEL FROM AGE GROUP TO AGE GROUP

Those categories showing consistent ratings from grade to grade will be discussed in terms of both sexes, of boys, and of girls.

Both sexes.—The following categories drew high and consistent ratings from boys and girls in all grades: newspaper comics and sports (although the percentage score for comics fell off somewhat while the score for sports rose); comedy in the

movies and on the radio; humor in magazines and books; mystery and detective in radio, movies and books. Sports and movie adventure were more popular with boys than girls. Humor in magazines and books was more popular with girls than boys.

A consistent secondary interest was shown in horror movies; crime, cartoons, history and historical fiction and mystery in magazines; disaster, religion, and art-literature-music in types of news.

Consistently low in the preferences of both sexes were advertisements, book reviews, public affairs, and gossip columns, in the newspaper; letters to the editor in newspapers and magazines; business and industry in all media; news, religious programs, talks and columnists on the radio; social problems in all media; serious drama in the movies; farm, literary, and news magazines; travel articles in magazines; labor problems, military affairs, and politics in types of news.

Boys.—The following categories rated consistently high with the boys in all grades: war in movies and books; adventure stories in magazines; and sports magazines. All were more popular with boys than girls. The boys showed a secondary interest in humorous books, adventure and movies magazines, and history in books.

Low on the boys preferred lists were columns on love and personal problems in the newspaper; love and romance in movies, books, and magazines; fantasy, and religion in the movies; magazines for girls, women and young women; home service and romance magazines; art and music in books and magazines; Hollywood in magazines; biography in books; and news of love and marriage problems.

Girls.—The following categories rated consistently high with the girls in all grades: romance, movie, and girls' magazines; humorous books; magazine stories about the movies and Hollywood; and crime news. All were more popular with girls than boys, except crime news.

The girls showed a secondary interest in columns on love and personal problems in the newspaper; religion and fantasy in the movies; home service and picture magazines; books on sports, religion and art or music; biography; art and music in magazines; science and invention in books and types of news; and love and marriage problems in types of news. Historical movies; adven-

ture, science, boys', men's, and science magazines; and war books were low among the girls' preferences at all grade levels.

III. THREE HYPOTHETICAL PATTERNS OF INTEREST

In order to get a more generalized picture of interest tendencies among young people, three hypothetical interest patterns were established, under which specific subject-matter categories could be classified. It was desired to uncover the extent to which these patterns characterized any given grade-sex group.

The three hypothetical patterns were (a) Adventure and Violence; (b) Love, Private Life and Glamor; and (c) 'Educational' content. All categories (in each medium) which were subsumable under these headings were grouped together for analysis.

Mean preference ratings for each pattern were calculated for each grade-sex group by dividing the sum of the ratings for all applicable categories in each and every medium, by the total number of occurrences of the categories in each and every medium. There were twenty-two total occurrences in the adventure and violence classification; thirteen occurrences in the case of love, private life and glamor; and thirty-three occurrences in the case of 'educational' content. These ratings were subject to the characteristic disadvantage of means in that they could be skewed by extreme cases; however, there were few extreme cases in the distributions concerned. Figure I shows graphically the mean ratings of the three hypothetical patterns of interest by grade and sex.

The results indicated marked sex differences in the preferences for adventure and violence, and love, private life and glamor. The boys' ratings exceeded the girls' in every violence and adventure category. The girls' ratings exceeded the boys' in every love, private life and glamor category. Neither sex indicated any substantial preference for 'educational' content.

There were also clearcut differences by grade in the case of adventure and violence, where the relative preference of both boys and girls declined steadily as the grade level rose. Differences by grade in the preference shown love, private life and glamor were not clear cut.

It should be emphasized that most of the individual categories

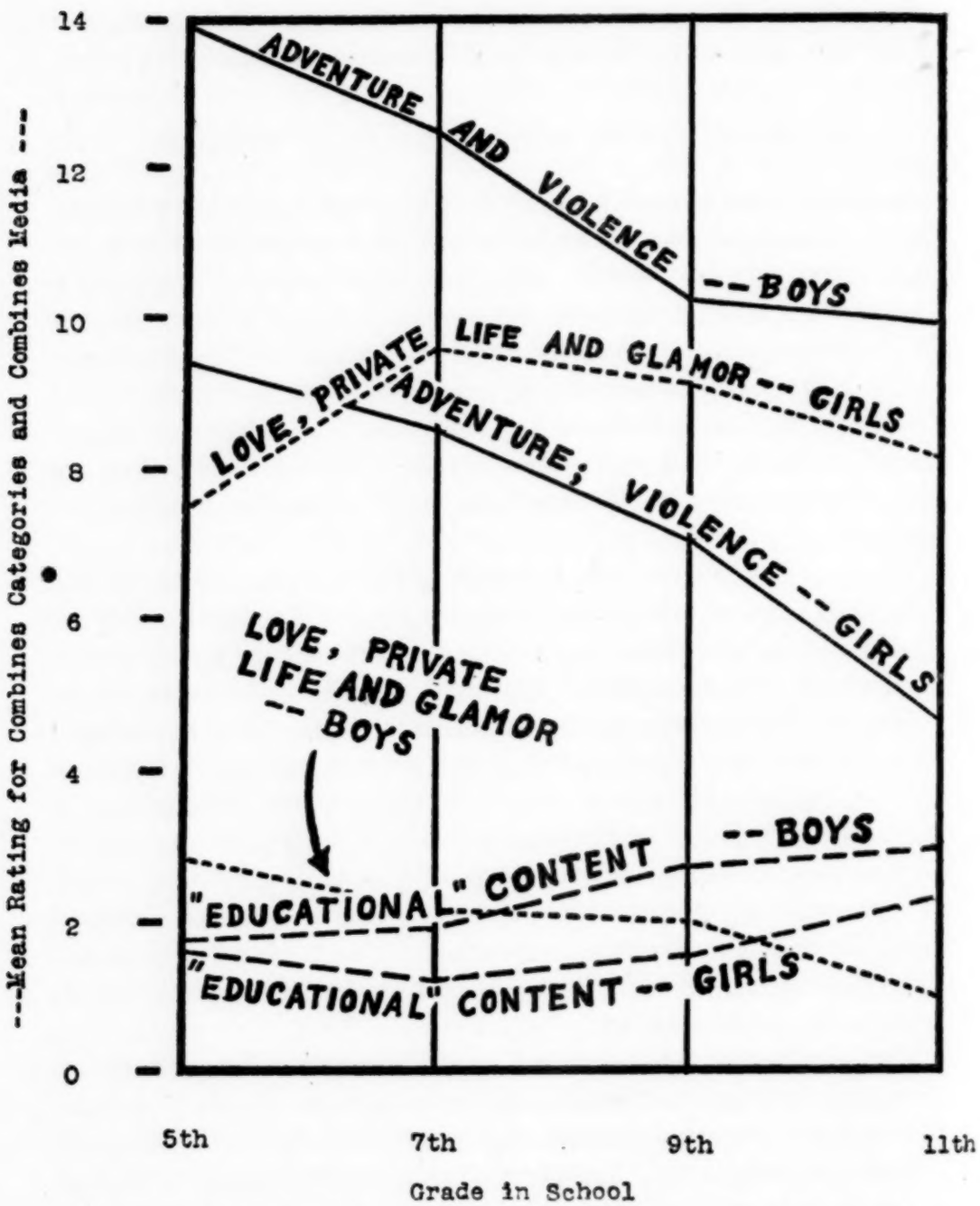

FIGURE I. Comparison of Three Hypothetical Patterns of Interest, by Preference Level, Sex and Grade.

making up each pattern occurred in more than one medium; for example, mystery occurred in radio, movies, magazines and books. Accordingly, the interests discussed in this section reflect subject-matter preferences in all media combined.

Examination of individual categories revealed certain distinct

trends related to sex and grade differences. Within the adventure and violence pattern, interest in westerns dropped sharply with increasing age levels. Fifth-grade boys and girls gave westerns a far higher rating than any other category; yet eleventh-grade boys and girls gave westerns a relatively low rating. The same tendency could be observed in ratings of radio adventure serials. The ratings for mystery, crime, disaster, war, and adventure other than radio serials, remained at about the same level from grade to grade. Also, interest ratings for these five categories were roughly of the same magnitude within grades. These tendencies applied to both sexes, although in all cases boys gave higher ratings to adventure and violence than girls. Both sexes rated horror lowest among all categories classifiable as adventure and violence.

All individual categories classified under love, private life and glamor were given relatively inconsequential ratings by boys. However, some interest in movies and Hollywood was shown by boys, especially in the fifth and seventh grades. Girls gave high ratings to love and romance, and movies and Hollywood. A tendency for the love and romance ratings to rise as the girls' grade level advanced was noted Conversely, interest in movies and Hollywood appeared to drop progressively as the girls' grade level advanced. Considerably less female interest was shown in the remaining categories classified under love, private life and glamor; that is, radio love serials, news on love and marriage problems, newspaper columns on love and personal problems, radio columnists, and newspaper gossip columnists.

Eighteen individual categories were classified under 'educational' content. With a few exceptions, these categories were given low ratings by both sexes. Relatively high ratings were given science and international and foreign news by ninth- and eleventh-grade boys. History elicited a moderate preference rating from boys, regardless of grade, and from eleventh-grade girls. Ninth- and eleventh-grade girls also ranked international and foreign news notably higher than other 'educational' content. Fifth-grade girls showed a substantial preference for art, literature and music. These exceptions of course, were not reflected in the mean ratings shown in Figure I, because of the generally low ratings of all other categories under 'educational' content.

STRENGTH OF EACH PATTERN BY MEDIUM

An indication was desired of the consistency of interest in the three hypothetical patterns from medium to medium. Accordingly, the preference ratings of all categories applying to each pattern within each medium were averaged. For example: In order to obtain the adventure and violence rating for radio, the preferences for the three pertinent radio categories were averaged; that is, adventure serials, mystery plays, and western shows. For movies, there were five pertinent categories—adventure, horror, mystery, war, and westerns.

A considerable degree of consistency of interest from medium to medium was observed for certain grade-sex groups. Adventure and violence in all media appealed strongly to fifth- and seventh-grade boys. Love, private life and glamor in all media but radio appealed strongly to seventh- and ninth-grade girls. 'Educational' content in all media, except magazines, received a low preference rating among boys of all ages. The somewhat higher male interest in 'educational' magazines, was due to the substantial popularity of science and mechanics. The girls' ratings of 'educational' content were uniformly low in all media. On the other hand, marked media differences in the appeal of the hypothetical patterns to certain grade-sex groups were observed.

Analysis of differences in the adventure and violence classification revealed that the older the student the less consistency from medium to medium. Adventure and violence in radio and magazines was much less popular among eleventh-grade boys than fifth-grade boys, although male interest in movie and book adventure and violence continued at a high level in all grades. Girls' preferences for this pattern followed the same general tendency; that is, for media differences to develop in the upper grades. Little interest on the part of either sex was shown in adventure magazines. This may have been due to the relative vagueness of the category.

Girls' ratings of the love, private life and glamor categories revealed a tendency for magazine stories and features and movies to become more popular relative to the other media from the fifth through the eleventh grade. Conversely, the relative popularity, for girls, of magazines devoted entirely to love and glamor appeared to decrease sharply after reaching a peak in

the seventh grade. Radio categories in the love, private life and glamor area had no significant appeal for either sex at any grade level.

Under love, private life and glamor, radio and books were the weakest media. The two radio categories were love serials and radio columnists. The book category was love and romance.

The male preference for "educational" content in magazines, as against other media, was due in large part to an interest in magazines on scientific and mechanical subjects.

THIRD GRADE PREFERENCES

Questions used in the third-grade interviews were largely of the free answer type; no check lists were used. Hence, the results were analyzed separately from those obtained in the higher grades.

Third-grade tastes, as revealed by this study, were focused mainly on the comics, westerns, animal themes, the world of make believe, adventure, comedy and humor, and mystery.

By medium, the prevailing choices of both boys and girls were comics in the newspaper; westerns on the radio and in the movies; fairy tales and fantasy in books; adventure serials on the radio; comedy and humor in movies and magazines; animal themes in magazines and books, and mystery in radio fare, movies and magazines.

Boys, more than girls, favored western movies and books; and mystery on the radio. Girls, more than boys, favored comedy and humor in movies and books; and animal themes in books.

Third-graders indicated little preference for love, private life and glamor, although a considerable number expressed their liking for the movie and radio sections of the newspaper. They also show little interest in informative, non-entertaining content in any medium.

AN ANALYSIS OF A PROCEDURE USED IN THE TEACHING AND RATING OF ORAL EXPRESSION

FRANCIS J. DI VESTA and JAMES H. L. ROACH

Air University
Maxwell Air Force Base, Alabama

The Air University's Air Command and Staff School, like other schools of higher learning, has been faced with the problem of evaluating various types of seminar performance not easily measured by objective tests. Among these is the problem of rating students' performance in oral presentation. Several methods have been tried for such ratings. New methods are being tried. The purpose of the present study is to evaluate the results of one general procedure used in teaching and rating oral expression at the Air Command and Staff School. This material is presented because the unique character of the course and rating system may be of benefit to those engaged in similar educational activity.

DESCRIPTION OF THE COURSE

The course in oral expression has been established to develop, in Air Force officers, the ability to organize speaking material and to present it logically and forcefully. This objective arises from the very practical needs of Air Force officers to speak before civilian organizations, to participate in Air Force conferences or joint conferences with representatives of the other services and the departments and agencies of the Federal or State governments, to brief their commanding officers, and to participate in staff meetings. Because the length of the course is only fifteen class hours, the instruction is oriented around formal oral presentation on the assumption that speaking before large groups would be the most difficult type of requirement to be met in the field. Only the elementary fundamentals of oral expression are taught. Students are not expected to be finished conference members or speakers at the end of the course.

The instruction includes three one-hour lectures on essential steps in preparing a speech, divisions and organization of a speech, and effective speaker projection. The remaining class periods are devoted to presentations by students. Each student is required to make four oral presentations. The first requirement is a five-minute talk on some phase of preparing a speech. The second requirement is a ten-minute talk on a subject of the student's choice, but of military significance. The third requirement is a five-minute impromptu presentation on the student's choice of one of four subjects within the broad areas of operations, personnel, logistics, and world events. This is a test of the student's ability to delimit a broad subject and to speak spontaneously on that subject. The fourth requirement is a prepared fifteen-minute talk wherein the student regards himself as a numbered Air Force staff officer required to give a speech on some phase of the military establishment to a civilian audience.

The total enrollment of four hundred sixty-five students is divided into small groups of about ten students each for the oral presentations. An instructor is assigned to each group. Both instructor and students rate each oral presentation on a rating scale described later in this report. All students are required to make the first three presentations. Those students in the top twenty per cent of the class are exempted from making the last presentation. Each rating provides a basis for grading the student and for constructive criticism.

DESCRIPTION OF THE RATING SCALE

The rating scale, devised by the Evaluation Branch of the Air University's Educational Advisory Staff, was made up of twelve items rated on a scale of one to five. Items were selected on the basis of characteristics which, it was believed, could be observed and remembered during the time the speaker was on the platform and which could be used as a guide for the speaker while preparing his speech. The number of items was limited to reduce rating time as much as possible and still provide sufficient coverage of qualities making for good performance.

These considerations were necessary since student-raters were being used and a number of ratings were made during each class period. The rating scale follows:

EVALUATION FORM FOR ORAL PRESENTATIONS

I. *Platform Presence.*—Does speaker appear poised and purposeful? Is he well groomed?

1	2	3	4	5
Ill-at-ease and untidy appearance detracts from platform presence.		Slightly self-conscious; presentable; acceptable platform presence.		Poised and immaculate; excellent platform presence.

II. *Introduction.*—Is introduction effective in getting attention, and audience acceptance? Is objective understood? Did he establish importance of subject and lead smoothly into it?

1	2	3	4	5
Inadequate introduction; failed to accomplish one or more of the purposes of an introduction.		Satisfactory introduction, fairly complete; adequate but not effective.		Complete and very effective introduction.

III. *Actions.*—Do posture, body movements, and gestures increase effectiveness of presentation?

1	2	3	4	5
Unnatural, awkward actions; annoying mannerisms definitely distract from presentation.		Fairly natural actions; annoying mannerisms slightly distracting.		Natural actions greatly aid presentation; no distracting mannerisms.

IV. *Eye Contact.*—Does speaker look at each member of the audience?

1	2	3	4	5
Eye contact detractive; tended not to look at audience.		Eye contact fairly effective; tended to look at audience		Eye contact very effective, well distributed, constantly looked at audience.

V. *Voice.*—Is voice pleasing? Can you easily and clearly hear him? Is he fluent and forceful?

1	2	3	4	5
Voice boring, irritating; indistinct, hesitant delivery.		Voice pleasing, fairly clear; matter of fact delivery.		Voice very pleasing and clear; eloquent, stimulating delivery.

VI. *Phraseology.*—Does he pronounce and use words correctly? Are sentences short, clear and complete? Does he use terms clear to audience?

1	2	3	4	5
Annoying grammatical errors. Poor choice of words. Does not define terms; difficult to follow.		Minor grammatical errors. Satisfactory choice of words; meaning generally clear.		Excellent grammar and sentence structure; effective choice of words; easily understood.

VII. *Body.*—Is plan of presentation clear to audience? Are major points effective and properly supported by minor points and details?

1	2	3	4	5
Plan not clear, too many or too few points inadequately supported.		Plan fairly clear; fairly adequate major points, moderately supported.		Plan very clear; proper number of major points, well supported.

VIII. *Transition.*—Does speech move smoothly and effectively from each part or point to following part or point? Does it have continuity?

1	2	3	4	5
Disjointed transition; parts not fitted together; lacks continuity.		Fair transition between parts; fair continuity.		Smooth transition; parts fit naturally together; excellent continuity.

IX. *Knowledge of Subject.*—Did speaker have sufficient background of facts on subject? Did he understand implications of material? Were his conclusions sound? Did he adequately cover and emphasize important aspects of subject?

1	2	3	4	5
Lacked essential knowledge of subject and its implication; conclusions faulty; failed to cover important aspects.		Fair knowledge of subject; conclusions generally adequate; partially covered important aspects.		Knew subject thoroughly; drew sound conclusions. Effectively covered important aspects.

X. *Conclusion.*—Does conclusion effectively end speech? Does it have punch? Is it consistent with purpose?

1	2	3	4	5
Weak, inadequate conclusion; new material introduced. Leaves audience up in air, not consistent with purpose.		Fairly effective ending, fairly consistent with purpose; lacks punch.		Concise, forceful ending; drives home purpose.

XI. *Preparation.*—Was speech extemporaneous (not memorized), and was it delivered without excessive use of props such as manuscript, notes or lectern? Was it within time limits and was proper time apportioned to developing each point or part? Were visual aids effectively used?

1	2	3	4	5
Indicated inadequate preparation by excessive use of props; poorly apportioned timing with speech too long or too short; faulty use of visual aids.		Indicated fair preparation by moderate use of props; fairly good timing; acceptable use of visual aids.		Indicated thorough preparation by no reliance on props; excellent timing and effective use of visual aids.

XII. *Interest.*—Did speaker establish and maintain audience interest in his subject and in himself as a speaker? Was audience reaction to speaker and subject friendly?

1	2	3	4	5
Dull and boring speaker, no audience interest; audience reaction indifferent or unfriendly.		Moderate audience interest; audience reaction fairly cordial.		Intense audience interest; audience reaction very friendly.

The grade for any one oral presentation was composed of the average of all students' ratings on this form added to the instructor's rating. Ratings are corrected for rater bias by simply adding or subtracting the rater's deviation from the mean of all ratings. Instructor ratings and students' ratings are corrected independently.

An indication of the reliability of ratings made on this scale was determined by finding the correlation between the average of student ratings and instructor ratings on each student's presentation in the second oral requirement (a ten-minute prepared speech on a subject of military significance). The cor-

relation between four hundred sixty-five scores paired in this way was .63 ± .02. Although this correlation is not as high as is generally desired for reliability coefficients, it does represent a significant relationship of better than moderate magnitude.

Further information about the rating scale is presented in Table I where intercorrelations of ratings on each characteristic with ratings on every other characteristic and the total score are shown. The item numbers in this table correspond to the item numbers in the rating scale.

TABLE I.—INTERCORRELATIONS BETWEEN SCORES ON EACH OF THE TWELVE FACTORS OF THE EVALUATION FORM FOR ORAL PRESENTATION AND THE TOTAL SCORE*
(N = 465)

Item Number†	Rating Scale Item Number*												Total
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
I		.46	.57	.47	.55	.48	.42	.61	.44	.44	.35	.53	.73
II			.33	.29	.37	.35	.44	.42	.42	.40	.29	.48	.62
III				.52	.42	.36	.30	.38	.41	.31	.26	.42	.62
IV					.38	.34	.11	.41	.35	.31	.26	.44	.61
V						.47	.36	.42	.38	.38	.33	.46	.65
VI							.42	.51	.47	.45	.37	.46	.68
VII								.68	.68	.50	.50	.63	.75
VIII									.68	.50	.54	.58	.79
IX										.50	.53	.62	.77
X											.56	.56	.70
XI												.46	.65
XII													.78

* Data is from the instructor ratings on Oral Presentation III only.

† Item numbers correspond to numbers opposite the rating scale items shown in the text.

These intercorrelations indicate the existence of a positive relationship among all items. Relatively higher relationships exist between each item score and the total score. These relationships are desirable because they indicate that, in general, each item evidently makes some unique contribution to the total score. None of the items correlates so highly with the total score that it could replace the entire scale. Items VII, VIII,

IX, and XII have relatively higher intercorrelations than other items. This is to be expected, since they are also related in the practical situation. Each of these four factors is essential to preparation of the subject-matter content of the speech. These results have not been factorized, however, to determine the extent to which the assumed factors or other factors actually are present.

RESULTS

The effects of the oral presentation program are illustrated in the data obtained from the rating scales.

Table II shows the mean scores and standard deviation for each of the oral presentations. The scores shown are a composite of the average of nine student ratings plus one instructor rating. The total possible score is thereby 120. From this table it will be recognized that the rating scale reflected an increase in scores as the course progressed. Differences in mean scores between any one speech and each of the other speeches are significant at the one per cent level.

TABLE II.—MEANS AND STANDARD DEVIATIONS OF THE ORAL PRESENTATION SCORES*

Speech Number	Number	Mean	SD
I	460	85.8	10.75
II	460	90.9	9.45
III	460	89.1	9.82
IV	360	92.5	8.16

* Differences between means on Oral Presentations I and II, I and IV, II and III, and III and IV are significant.

The lower mean score for Speech III may be interpreted as evidence of the validity of the scale. Speech III is the five-minute impromptu speech with only ten-minute preparation. On this basis it would be expected that the individual's work would be poorer than on a prepared presentation. This interpretation is reflected in the lower mean score for the third speech than for the second or fourth speech.

Intercorrelations between each speech and every other speech are shown in Table III. These correlations show a substantial relationship between student performance early in the course and

later performance. Lower correlations between scores on Speech IV and scores on other speeches are undoubtedly a result of the increased homogeneity of the class (the top twenty per cent on three earlier speeches were excused from making the fourth speech).

TABLE III.—INTERCORRELATIONS BETWEEN EACH OF THE ORAL PRESENTATIONS AND FINAL ORAL PRESENTATION SCORES CLASS 49-B, AC&SS

Oral Presentation	Oral Presentation				Final Score N = 460
	I N = 460	II N = 460	III N = 460	IV N = 360	
I		.67	.58	.40	.85
II			.69	.49	.89
III				.42	.84
IV					.71
Final Score					

A study was made of the progress of those individuals in the highest ten per cent of the class (N = 46) and those in the lowest ten per cent of the class (N = 46) selected on the first presentation. The averages for these groups are shown in Table IV. The significance of the differences between the scores for each speech for each of these groups is shown in Table V.

Although there is a drop of five points from average of Speech I to Speech III for the highest ten per cent, the difference is significant at slightly more than the five per cent level of con-

TABLE IV.—MEANS OF SCORES FOR EACH ORAL PRESENTATION MADE BY THE HIGHEST AND LOWEST 10 PER CENT OF STUDENTS ON ORAL PRESENTATION NO. I

Speech	Highest 10% on Speech I	Lowest 10% on Speech I
I	105.2	67.4
II	101.3	79.2
III	99.9	80.8
IV		86.7

fidence. None of the other differences in speech scores for this group is significant.

The lowest ten per cent of the class on Oral Presentation I made gains significant at least at the two per cent level of confidence except for the progress shown from Oral Presentation II to III. Again, this is probably accounted for by the fact that Oral Presentation III was an impromptu requirement.

TABLE V.—SIGNIFICANCE OF DIFFERENCES BETWEEN ORAL PRESENTATION AVERAGES MADE BY THE HIGHEST AND LOWEST 10 PER CENT OF THE STUDENTS SELECTED ON THE BASIS OF SPEECH I SCORES
(N = 46 for each group)

Differences Between Speeches:	Highest 10%			Lowest 10%		
	DF	X^2	P	DF	X^2	P
1 and 2	2	5.21	>.05	2	13.21	<.01
1 and 3	2	3.28	<.20	2	19.20	<.001
1 and 4				1	32.76	<.001
2 and 3	2	.07	.98	2	1.66	<.50
2 and 4				2	14.28	<.001
3 and 4				3	9.63	>.02

RELATIONSHIP BETWEEN PERFORMANCE IN ORAL EXPRESSION AND OTHER FACTORS

Correlations were found between the final oral expression grades and several other factors on which scores were available (see Table VI). The data studied were as follows: *General Survey Test*—a two-part test yielding a score on student achievement in Air Force background subject matter, another score for achievement in mathematics, science, and English and a total score; the *American Council on Education Psychological Examination's* Q, L, and total score; the student officer's rank; the student officer's period of active service; the student officer's age; consolidated scores of fifteen class examinations; and grades on the final school problem, which is a simulated problem situation solved by a staff of sixteen student officers.

TABLE VI.—CORRELATION OF FINAL ORAL EXPRESSION GRADES AND TWELVE FACTORS
(N = 465)

Factor	Correlation with Oral Expression Grades
General Survey Test	
Military subjects	.31
Basic	.26
Total	.33
ACE Psychological Examination	
Q score (quantitative)	.22
L score (language)	.30
Total	.29
Rank	.00
Months active service	.10
Age	.06
Class examinations (15 consolidated scores)	.47
Final school problem (staff problem)	.22
Final class standing (including oral expression)	.50

From these correlations, very little relationship is shown to exist between oral expression and rank, age, and months of active service. The oral expression grades have a higher relationship with general intelligence, as measured by the *ACE Psychological Examination* ($r = .30$), staff action ($r = .22$), and military achievement scores ($r = .31$). A still higher relationship is noted between oral expression and classroom achievement on written examinations ($r = .47$). It is interesting to note that the oral expression grades correlate more highly with the *American Council on Education Psychological Examination's* L score ($r = .30$) than with the Q score ($r = .22$). The difference between these correlations is, via Fisher's Z transformation, significant at the one per cent level of confidence.

SUMMARY

The oral expression course at the Air University is established for the purpose of imparting basic principles in the preparation and delivery of speeches. Practice is given through four speech

requirements. Appraisal of student performance for purposes of evaluation and guidance is made by student and instructor ratings of each speech.

Results of the analysis of scores on the evaluation form for oral presentation indicate that (1) raters are differentiating between presentations; (2) growth and development in oral presentation takes place throughout the course; (3) no progress in oral presentation was reflected in the evaluation scale for those already qualified at this level of oral expression; (4) those who were originally very poorly qualified in oral expression developed to a point where their average score on the final speech requirement exceeded the entire class average on the first oral presentation. One might conclude that, as a group, the Air University students improve in oral expression throughout the period of instruction and, furthermore, that most improvement is made by those individuals who were poorest in oral expression before attendance at the Air University.

THE EFFECT OF ATTENTION-GAINING DEVICES ON FILM-MEDIATED LEARNING*

D. MORGAN NEU

The Pennsylvania State College

INTRODUCTION

Film-makers have employed a wide variety of techniques to achieve emphasis and increase the teaching effectiveness of their products. In general, the devices employed to direct attention towards particular points of the content presented may be divided into two broad classes: 'relevant devices' and 'irrelevant devices.' 'Relevant devices' include any emphasis technique that is related to, or operates upon, the specific point to be emphasized. For example, an ultra-close-up of part of a gun brings attention specifically and forcefully to that part of the gun. 'Irrelevant devices' include any emphasis techniques that call attention to the screen but are otherwise unrelated to the film content. For example, interrupting a film on a technical subject with a shot of a bathing beauty is an instance of an 'irrelevant device.'

While current learning theories and the experiments that support them are not entirely in accord on the relative effectiveness of these two methods of ensuring the perception and learning of the salient points in a lesson, the consensus seems to be that, in the long run, 'relevant devices' should facilitate recall, while 'irrelevant devices' should act as distractors and inhibit recall. This inference is obviously at variance with the results hoped for by producers who insert sequences that attract attention to the screen without relating to a specific point of information. It is important, therefore, to determine experimentally the effects on learning of such different attention gaining techniques.

STATEMENT OF THE PROBLEM AND HYPOTHESES

The major purpose of this study was to determine the relative effectiveness of three levels of attention-gaining devices; (1)

* Summary of a dissertation completed in pursuance of the requirements for the Ph.D. degree at The Pennsylvania State College, 1950. The work described here was accomplished under the Instructional Film Research Program under Contract N6 onr-269, Task order VII, with the Special Devices Center, Office of Naval Research.

relevant devices, (2) no devices at all, (3) irrelevant devices. Since it is possible to insert devices in the sound track or the visuals, the rôle of the medium of presentation has also been studied. In order to maintain a clear cut comparison between the sound and visual media, sound and visual attention devices were each studied separately. Three hypotheses were formulated for testing as follows:

1) *Device relevancy.*—The hypothesis is that film-mediated learning will be facilitated by the use of relevant attention-gaining filmic devices, and inhibited by irrelevant attention-gaining devices. The hypothesis was tested by comparing versions of a film containing, respectively, relevant devices, no devices, and irrelevant devices.

2) *Device medium.*—The hypothesis is that film-mediated informational learning will be equally facilitated by visual and sound devices of the same relevance. The hypothesis was tested by comparing learning from visual device versions with learning from sound device versions. This comparison is made separately for relevant and irrelevant devices.

3) *Device recall.*—The hypothesis is that recall of the devices themselves is relatively independent of learning of the factual information content of the film.

EXPERIMENTAL DESIGN AND PROCEDURES

The Experimental Films.—To accomplish the comparisons required by the hypothesis, five versions of a film presenting an introduction to machine shop measuring instruments were produced. The five versions included the following:

1) Basic (identified as B in the tables): contains no experimental emphasis devices, but gives a clear, straightforward treatment of the subject.

2) Visual relevant (VR): contains attention-gaining devices in the visuals that are related to the points of content being emphasized.

3) Visual irrelevant (VI): contains attention-gaining devices in the visuals that are unrelated to the points of content being emphasized.

4) Sound relevant (SR): contains attention-gaining devices in the sound track that are related to the points being emphasized.

5) Sound irrelevant (SI): contains attention-gaining devices in the sound track that are unrelated to the points being emphasized.

Choice, number, and placement of the devices.—To ensure satisfactory comparability among the four 'device' versions, a panel of judges selected four sets of devices considered to be comparable in intensity and probable effectiveness in gaining attention.

It was decided that the devices selected should not add information to the film, but should be used solely for their possible attention-gaining value. Furthermore, it was decided to select attention-gaining devices that would represent various categories of stimulus conditions: movement and change, size, intensity, repetitive or prolonged contact, novelty and peculiarity, impressiveness, and sudden contrast.

A device was inserted at each of the same twenty-six most important points in each version. Table 1 lists the devices and their placement in relation to the informational content of the film. Thus, a point emphasized in one version by a visual relevant device (e.g., a spotlight) is also emphasized in the other three versions, by appropriate devices. It will be noted that some devices are used more than once.

THE TESTS

Two tests were constructed to provide measures of (1) learning of information and (2) recall of devices. The information test comprised one hundred four multiple-choice items, of which twenty-five provided picture-choices rather than verbal choices. This test had a high reliability (.94 by Kuder-Richardson Formula 20).

The device recall test included seventy-seven items, each of which briefly identified a device by a short statement. The subjects were called upon to check whether (1) the item was in the film, (2) the item was not in the film, or (3) they did not remember whether it was in or not.

THE POPULATIONS

The experiment was conducted with two separate military populations: (1) 1576 Army recruits at Fort Dix, New Jersey, (2) 1055 Navy recruits at The Great Lakes Naval Training Station, Great Lakes, Illinois.

TABLE 1.—DESCRIPTION OF ATTENTION-GAINING DEVICES

Basic Version	Attention-gaining Devices Added to Basic Version			
	Visual Relevant	Visual Irrelevant	Sound Relevant	Sound Irrelevant
1. Three metal blocks, rectangular, square, and round, are used to illustrate length, width, and depth.	White line appears on each block to emphasize length, width, and depth.	Concentric circles flash on screen as length, width, and depth are mentioned.	Long, short, and low tones emphasize length, width, and depth.	Auto horn sounds as length, width, and depth are mentioned.
2. The scale graduations on each end and both sides of the steel rule are shown one at a time.	Each scale graduation is spot lighted as it appears.	Train-crossing signals flash on before each graduation appears.	A female voice names each graduation as it appears.	A train whistle sounds before each appearance of the graduations.
3. A hook rule is used to measure from the shoulder of a metal block.	The rule appears to move by itself onto the shoulder.	Still shot of an athlete in action appears before and after rule is put on shoulder.	Before hook rule appears commentator says, "Now hear this. You are to remember. . . . "	Commentator loudly clears throat before discussing hook rule.
4. A slender flexible rule is held up for close-up shot.	End of rule is bent back and released which causes it to vibrate.	Portrait of pretty girl appears before flexible rule is discussed.	A 2nd male voice gives the name of the rule.	Sound of buzzing airplane before flexible rule is described.
5. The steel rule is shown so that the worn corners on one end are plainly visible.	Extreme close-up of this worn end.	The entire scene is tinted pink.	The commentary is read, slowly with emphasis.	Sound of applause in the background of entire scene.
6. The steel rule is wiped with an oiled rag several times.	Picture of oil can is flashed on the screen with the mentioning of oil.	Picture of cow's tail.	A squeaking sound is heard as oiled rag is rubbed across rule.	Fire siren starts when rule is rubbed with oiled rag.

7. The spring caliper is closed properly by squeezing the legs together.	Picture of hand squeezing the 1st and 2nd fingers of other hand together.	Concentric circles appear as caliper legs are closed.	Crunching sound as caliper legs are closed.	Sound of music as caliper legs are closed.
8. A group of adjustable precision instruments (micrometer and vernier caliper) are shown on workbench.	As each is mentioned, one micrometer and one vernier caliper turns around by itself.	Picture of a hammer hitting a nail as each instrument is mentioned.	The word micrometer, said with a rising inflexion, is followed by an ascending whistle. After the word vernier caliper is said with falling inflexion, a descending whistle follows.	After each word is said, a gurgling sound-like pouring liquid from a jug is made.
9. Two parts of the micrometer, the barrel and thimble, are discussed.	Finger points to the thimble.	Shot of a smoke stack before thimble is mentioned.	A 3rd male voice names the thimble.	Sound of squeaking door before thimble is mentioned.
10. A piece of paper is drawn between the micrometer's measuring surfaces to clean them.	Close-up of paper being drawn through.	Picture of turkey's head appears.	Resinous twang is heard as paper is pulled through.	Sound of cuckoo clock as paper is pulled through.
11. Micrometer is adjusted to show that graduations should line up to zero when measuring surfaces are closed.	Spot-light on measuring surfaces and on scale graduations.	Shot of close-up of two eyes looking at camera.	2nd male voice reads commentary.	Sound of sawing wood.
12. Micrometer is twirled to show "wrong way" of obtaining rapid adjustment.	Micrometer is twirled very rapidly.	Picture of man getting hit in the jaw.	Ratchet sound while micrometer is twirled.	Sound of miscellaneous objects falling down and breaking. (Fibber McGee's closet.)

TABLE 1.—*Continued*

Basic Version	Attention-gaining Devices Added to Basic Version			
	Visual Relevant	Visual Irrelevant	Sound Relevant	Sound Irrelevant
13. Micrometer approaches moving object in lathe as if to measure it.	Restraining hand stops the micrometer.	Scene goes in and out of focus.	Commentator says, "Now watch this closely."	Sound of jazz music.
14. A micrometer which measures from three to four inches is called a four-inch micrometer.	Six-inch steel rule is placed along side micrometer.	Concentric half-circles appear on the screen.	Female voice reads commentary.	Sound of water splash and gurgling effect.
15. The vernier caliper is used to measure a part in a lathe.	Part of scene at unusual angle from under the operator's arm.	Picture of profile of man' face.	Background sound of lathe running.	Sound of baby crying.
16. Snap gage is shown in comparison with a micrometer in background.	Snap gage is moved to extreme close-up in foreground with micrometer in background.	Pitcher throws ball at camera.	3rd male voice reads commentary.	Sound of breaking glass.
17. The three part of the snap gage (anvil, go button and not-go button are described as gage is shown on workbench.	Finger points to each part as it is mentioned.	Concentric half circles appear before each part is mentioned.	Sound of anvil when anvil is mentioned, sound of "go" police whistle for go button, and "stop" whistle for not-go button.	Sound of horse whinnying as each part is mentioned.

18. A part is inserted in a ring gage several times.	Zoom to close-up.	Picture of man lying on ground, feet nearest camera.	Ascending whistle when part is inserted in gage.	Pistol shot when part is inserted in gage.
19. Snap gage is inspected for accuracy with gage blocks.	Circular mask to concentrate attention on gage blocks.	Picture of locomotive coming at camera.	2nd male voice reads commentary.	Sound of fog horn.
20. Straight plug gages and taper plug gages are shown on workbench.	All but one straight gage disappear and reappear. All but one taper gage disappear and reappear.	Picture of baby as each group of gages is mentioned.	After "straight plug gages" said with rising inflexion, an ascending whistle. After "taper plug gages" a descending whistle.	Sound of drum roll as each group is mentioned.
21. Shot of hole in metal part. Hole contains many burrs.	Burrs gradually disappear by a dissolve.	Picture of airplane diving through sky.	Sound of filing metal as burrs disappear.	Sound of telephone as burrs disappear.
22. Measuring instruments are being checked in inspection laboratory.	Shot of large white check mark.	Picture of hammer hitting nail.	Voice says, "O. K."	Sound of hammer hitting nail.
23. Gage blocks are used to check setting of fixed gage.	Unusual angles: close up of gage with inspector's eye appearing through curve of gage; gage is seen in silhouette.	Still picture of athlete in action.	Female commentary.	Sound of rooster crowing.
24. Two gage blocks are "wrung" together.	Circular mask makes operation appear as through a hole.	Picture of gun being fired at camera.	Sound of woman's scream as blocks are "wrung."	Sound of coughing as blocks are "wrung."
25. Shot of projection comparator.	Hand points to the comparator.	Picture of girl's legs.	2nd male voice reads commentary.	Sound of "gibberish" (Donald Duck).
26. Shot of the interference bands produced by the light wave instrument.	Extreme close-up of the interference bands.	Picture of boy and girl kissing.	Sound of voice counting "1, 2, 3, 4, 5."	Sound of kiss and sigh.

PROCEDURES

At each testing center, the following procedures were employed: intact companies of recruits were divided at random into six comparable groups. Five of these groups were each shown one of the film versions (B, VR, VI, SR, or SI) and, immediately after the film, given the information tests and the device recall test. The sixth (control) group was given the information test without having seen the film. Test responses were recorded on IBM answer sheets, and machine-scored.

STATISTICAL TREATMENT OF THE DATA

For the men in each population both a measure of general intelligence and a measure of mechanical aptitude were available. Mean scores for the information test for the six treatments were calculated, and these means were adjusted, by an analysis of co-variance, to take into account small differences in initial ability as measured by the general intelligence and mechanical aptitude tests. The adjusted means therefore reflect the performance to be expected of groups having equal mean general intelligence scores and equal mean mechanical aptitude scores.

To determine the relationship between the information test and the device test, scores on these two tests were correlated.

RESULTS

Comparison of the Film Versions

Information Test Results.—In Table 2 the mean scores on the information test are given for the six groups of enlisted men in the Army and in the Navy. The adjusted means are given in parentheses. Table 3 summarizes the differences among the adjusted information test mean scores.

The findings can be summarized for both the Army and Navy populations as follows:

1) For all film groups substantial learning resulted from viewing the film. In each case, the film group mean score was between one and two standard deviations higher than the mean score of the comparable control group that saw no film.

2) The 'sound irrelevant' group had a lower mean score, in each population, than any other film group. It is to be noted, however, that the visual irrelevant version was the most effective

TABLE 2.—RAW MEAN SCORES AND ADJUSTED MEANS OF NAVY AND ARMY POPULATIONS ON INFORMATION TEST AFTER VARIOUS FILM TREATMENTS

Treatment	Navy			Army		
	N	Mean	SD	N	Mean	SD
Visual Relevant (VR)	136	67.63 (69.55)*	15.18	248	47.84 (48.02)	17.36
Visual Irrelevant (VI)	148	65.45 (66.62)	16.67	229	49.83 (49.24)	17.26
Sound Relevant (SR)	135	70.64 (70.68)	14.19	226	48.12 (49.00)	16.31
Sound Irrelevant (SI)	155	66.67 (65.11)	15.17	242	46.35 (46.31)	17.71
Basic (B)	331	69.80 (69.52)	15.87	367	52.55 (52.09)	18.80
Control (C)	150	40.57 (39.88)	10.40	264	30.04 (30.41)	9.74

* Adjusted means are in parentheses.

TABLE 3.—DIFFERENCES AMONG ADJUSTED MEAN SCORES FOR INFORMATION TEST[a]

	VR	VI	SR	SI	B
			Navy		
VI	−2.93*	—	—	—	—
SR	+1.13	+4.06**	—	—	—
SI	−4.44***	−1.51	−5.57***	—	—
B	−.03	+2.90*	−1.16	+4.41***	—
C	−29.67***	−26.74***	−30.80***	−25.23***	−29.64***
			Army		
VI	+1.22	—	—	—	—
SR	+.98	−.24	—	—	—
SI	−1.71	−2.93*	−2.69*	—	—
B	+4.07***	+2.85**	+3.09**	+5.77***	—
C	−17.61***	−18.83***	−18.59***	−15.90***	−21.68***

[a] In this table, the mean for the group listed on the top is subtracted from that of the group listed on the left.

* Significant at the 5 per cent level of confidence.

** Significant at the 1 per cent level of confidence.

*** Significant at the 0.1 per cent level of confidence.

of all the versions containing attention devices for the Army population.

3) The remainder of the inter-film differences were inconsistent for the two populations, and generally insignificant. In the Army, the basic version was consistently more effective than any device version, while in the Navy it was about equal to the two relevant device versions.

4) No significant difference was noted as between the visual devices and the sound devices versions, combined without respect to relevance of devices. For the Navy, the weighted mean of the two visual devices versions was 68.02, as compared with a weighted mean of 67.70 for the sound devices versions. For the Army, the visual devices group mean, without respect to relevance of device, was 48.60, while the sound group mean was 47.61. (These are all adjusted means.)

Results of the Device Recognition Test

The recognition test items were divided into two groups for each film version: (1) those devices that were in the film (Test C_1) and (2) those devices that were not in the film (Test C_2).

In scoring responses for the C_1 test, which comprised the devices actually in the film, a plus one (+1) was given for each device reported as present, a zero (0) if the subject reported he did not remember, and a minus one (−1) if the device was reported not present. The reverse of the scoring of the C_1 test was used for the C_2 test which comprised all devices not in the version being tested. A plus one (+1) was given for each device reported as not present, a zero (0) if the subject did not remember, and a minus one (−1) for each device reported as present. These scores were totaled for each film version and the tests (C_1 and C_2) correlated with each other as well as with the film information test scores. Table 4 reports these correlations.

All the correlations between the film information test and the C_1 test, and the film information test and the C_2 test, are generally insignificant although positive, indicating that there was only a slight relationship between scores on the film test and the remembering of the devices. The correlation between the C_1 and C_2 tests is also low. In all cases except one, which was practically zero, the correlations were negative. Analysis of the data showed that there was a slight tendency for the individual to mark the 'not-present' devices as present in the

TABLE 4.—CORRELATIONS BETWEEN DEVICE RECOGNITION TESTS AND INFORMATION TEST: C_1 TEST-DEVICES IN THE VERSION, C_2 TEST-DEVICES NOT IN THE VERSION

Group	C_1 Test and Information Test		C_2 Test and Information Test		C_1 Test and C_2 Test	
	Navy	Army	Navy	Army	Navy	Army
Visual Relevant	.185	.191	.191	.330	−.295	−.079
Visual Irrelevant	.338	.299	.132	.344	.042	.053
Sound Relevant	.223	.092	.141	.203	−.431	−.225
Sound Irrelevant	.107	.075	.073	.150	−.416	−.282
Basic*			.214	.334		

* Since the Basic Version contained no devices there was no C_1 test score for this group.

film, if he marked the 'present' devices as present. In other words, there was a general tendency to indicate that devices were present, whether they were used or not.

An analysis of responses to the individual items included in the Device Recognition Test revealed two significant points. First, whereas only about half the items included in the VR, SR, and SI versions were recognized by seventy per cent or more of the groups seeing these versions, almost three quarters of the devices in the VI version were correctly recognized by seventy per cent or more of the groups seeing this version. In short, the visual irrelevant devices seemed to have called attention 'to themselves' to a greater extent than any other type of device did. Second, whereas a maximum of only seven devices were falsely recognized (i.e., they were not in the version but the subjects said they were) by forty per cent or more of the VR, VI, and SR version groups, fifteen items were falsely recognized by at least this proportion of the sound irrelevant group. It is suggested that the irrelevant sound devices tended to confuse the subjects so that they were uncertain about what they did hear.

SUMMARY OF THE RESULTS

The results may be summarized with respect to the three experimental hypotheses as follows:

Device relevancy.—There is no evidence that the insertion of relevant devices, as defined in this study, adds to the effectiveness of an informational film. There is some evidence that irrelevant sound devices detract from the teaching effectiveness of such a film. (For both populations, the sound irrelevant version yielded the lowest scores.) To a lesser degree there is evidence that any of the devices used may have the same effect. (For the Army populations the basic version—no devices—yielded the highest score.)

Device medium.—No significant difference was found between the effectiveness (or ineffectiveness) of visual and sound attention-gaining devices.

Device recognition.—Ability to recognize and remember which devices were used in a film version was practically independent of learning from the version.

CONCLUSIONS AND RECOMMENDATIONS

This study leads to the general conclusion that attention-gaining devices of the kinds inserted in these films do not add significantly to learning.

The following recommendations are suggested as a result of this investigation:

1) Where instruction is the principal aim, and cost a consideration, producers of training films should present the subject matter in a simple, straightforward way and avoid the use of such fancy and expensive devices as spotlighting, zooms, extreme magnification, and stop motion, to gain the learners' attention.

2) If it seems necessary, in an instructional film, to use devices to attract or direct the learners' attention, use a technique which will emphasize something already in the film—some special treatment of indigenous materials related to the subject content—rather than introduce extraneous or irrelevant materials.

The findings of this experiment are largely negative; however, that does not nullify the usefulness of the study. It is almost as important to discover that certain techniques are ineffective, as it is to learn that other techniques are effective. In general, this study corroborates the belief which many educators have held for a long time, that it is not necessary to have a fancy, expensive treatment to convey ideas by means of film, and that a simple straightforward treatment is often the most effective.

A COMPARISON OF READING ABILITY AND READABILITY

JOHN U. MICHAELIS and FRED T. TYLER

School of Education
University of California, Berkeley

Numerous attempts have been made to develop a satisfactory technique for determining the level of reading difficulty of printed materials. There are three of these which are given most extensive use at the present time; namely, those proposed by Lorge, Flesch, and Dale and Chall. Since their appearance many types of materials have been appraised by one or more of the formulae: textbooks and children's books;[20] advertising;[1] industrial publications;[13] pamphlets on international relationships;[10] newspaper materials;[18] corporate annual reports.[21] The formulae have been used by the American Library Association[4] and the United States Department of Agriculture Extension Service.[22]

Since these techniques essentially provide a type of 'measuring' instrument, they should be examined to determine their validity and reliability. The first study on reliability, i.e., analyst-to-analyst reliability, of the Flesch formula was reported by Haynes, Jenkins and Walker.[12] They conclude that "For practical purposes the Flesch formulas and the directions for their use are sufficiently objective to be used even by inexperienced analysts to obtain estimates of reading ease and human interest of written materials." As for validity, Flesch found that materials rated as more readable by his formula were also more readable according to readership surveys and to expert opinion. The studies which the present writers have located have relied upon this latter type of criterion of readability. The subjective nature of such a criterion raises the question of its reliability.

PURPOSE

For the present paper it is assumed that materials rated as readable at a certain grade level should be understandable to students at that level. Thus, if a readability formula indicates that certain material is located at the Grade IX level, then it is to be expected that children at the Grade IX reading level

should be able to read the material and then to answer questions on it with a high degree of accuracy.

The purpose of the present paper is twofold: (a) to determine the readability of certain United Nations publications by each of the three readability formulae mentioned above; (b) to determine the reading ability of a group of Grade XII students by means of some standardized reading test, and the accuracy with which these same students can answer questions based on the United Nations materials.

The need for such a study has been emphasized in numerous papers appearing during the last few years. Lostutter[18] concluded that the Flesch formula was best adapted for the measurement of the readability of newsprint. Lorge[17] has stated that each of the three formulae has a specific function to perform. He claims that his formula will give "evidence on the comprehension difficulty of the kind of reading pupils have to do in school," that the Flesch formula provides a measure of readability of magazine materials, and that the Dale-Chall formula "will make for adequate appraisal of health materials." Flesch, on the other hand, has pointed out that, "There is no theoretical justification for this neat pigeonholing of formulas; all three of those mentioned were based on exactly the same criterion; namely, the McCall-Crabs' Standard Test Lessons in Reading." Both Lorge and Flesch agree, however, that there is some degree of correlation between the formulae when applied to the same materials: "The statistical differences between their predictive values are insignificant." In a recent paper Michaelis[19] indicated that some U. N. materials were written on a difficult level according to teachers' judgments and data from readability formulae. He reported also that there was considerable disparity between the difficulty levels obtained by the different formulae.

TEST MATERIALS

The Iowa Silent Reading Test (Form Cm) was administered to forty-two boys and thirty-six girls in Grade XII at the Oakland Technical High School to determine the grade level at which these students were reading. The same students were also given an objective test based upon several U. N. publications which had been prepared for use in high-school social studies classes.

The U. N. test included four subtests: vocabulary, sentence comprehension, paragraph comprehension, and reading for details. The first part required the students to read a paragraph chosen from one of the U. N. pamphlets. The students were allowed three minutes to read and study the materials. At the end of that time the printed materials were collected and the students attempted to answer questions based on their reading. There were nine multiple-choice questions in Part I. The second part was a vocabulary test, again multiple-choice in nature. The test included twenty-two words and phrases, such as "preamble," "scope," "indigenous," "the framers of the Constitusion," and "ocean rendezvous." The third part consisted of a series of sentences, a phrase in each sentence being underlined. The students were required to indicate which of four choices could most reasonably be substituted for the underlined expression. The last part attempted to measure the students' understanding of short paragraphs. The multiple-choice questions were arranged to test the ability of the student to select the best title for the paragraph; to test his knowledge of word meanings; to test his ability to read for details. There was also a free expression question on each paragraph, but this item is not used in the present report. The reliability of the test of sixty-nine items, determined by the split-half method and stepped-up by the Spearman-Brown Prophecy formula, was .85.

FINDINGS

The readability of the U. N. materials was determined by each of the formula listed above. The Lorge formula considers average sentence length, ratio of prepositional phrases to number of words, and ratio of hard words to number of words in various samples of the materials being appraised. The Flesch formula for reading ease involves the determination of number of syllables per one hundred words and average sentence length. The Dale-Chall formula requires the calculation of average sentence length and the number of hard words. Readability by each formula was determined on the same sampling of U. N. publications. Five selections were chosen from each of five pamphlets according to the directions for the application of the formulae. Similarly the proper number of words was included in each sample. The computation of readability was done on two separate

occasions with practically identical results. Furthermore, five other samples were chosen from each of the five pamphlets. Readability was determined for each of these samples, the results checking with those obtained from the previous sampling. Also the U. N. test was prepared in the main from the materials that had been used to determine readability. The readability of the test materials and of the instructions for the test was also determined by each formula. The U. N. publications varied from seven to two hundred one pages in length and were believed to be representative of those available for use in the public schools.

TABLE I.—READABILITY OF UNITED NATIONS MATERIALS APPRAISED BY THREE FORMULAE

United Nations Material	Ratings According to Formulae					
	Dale-Chall		Flesch		Lorge	
	Raw Score	Grade Placement	Raw Score	Level of Difficulty	Raw Score	Grade Placement
Basic Facts	9.6226	13–15	28.6720	College graduate	8.6524	8+
Everyman's UNO	9.4727	13–15	24.2233	College graduate	8.3557	8+
FAO	8.2895	11–12	38.9930	13–16	7.5497	7+
Guide to UN Charter	9.8611	13–15	31.7220	13–16	8.1380	8+
How Peoples Work Together	9.5917	13–15	37.6400	13–16	7.9803	8
Average	9.3675	13–15	32.2500	13–16	8.1352	8+

Table I shows the readability of the five pamphlets computed by each of the three formulae discussed above. From the table it may be seen that the Lorge formula indicates that the materials should be readable at about the Grade VIII level. According to the Flesch formula the materials are rated at the college undergraduate and graduate levels. Readability as estimated by the Dale-Chall method is placed at Grades XI to XV for the different pamphlets. There is, then, considerable discrepancy between the results obtained from different readability formulae.

As has been mentioned, it has been assumed for this study that if materials are readable at a certain level, then students at that level of reading ability should be able to read the materials with a high degree of accuracy. Specifically this would mean that according to the Lorge formula children who read at about the Grade VIII level should be able to comprehend the U. N. materials. Table II shows the mean percentile rank on the Iowa Silent Reading Test of the seventy-eight subjects in Grade XII according to the manual of directions.

TABLE II.—MEAN STANDARD SCORES AND PERCENTILE RANKS OF SUBJECTS

Subtest	Mean Standard Scores	Percentile Ranks in Grade XII
1. Rate	171	41
1. Comprehension	203	92
2. Directed Reading	166	42
3. Poetry Comprehension	172	50
4. Word Meaning	179	56
5. Sentence Meaning	171	35
6. Paragraph Comprehension	176	58
7. Use of Index	177	63
7. Selection of Key Words	169	43

The mean standard score on the whole test was 174, which corresponds to a grade equivalent of 13.0, indicating that these students were on the average reading at about the Grade XIII level.

Before indicating the success of these subjects on the U. N. reading test, it is desirable to indicate the readability level of the test materials and of the instructions to the test. This is done in Table III.

The sampling of test materials was no more difficult than was the sampling of the U. N. publications; the instructions were readable at a slightly lower level than were the test materials. The U. N. test was slightly easier than were the pamphlets upon which it had been based because the readability score for the test included the estimate of the readability of the instructions. It now remains to indicate how well the students were

TABLE III.—READABILITY OF U. N. TEST AND OF TEST INSTRUCTIONS

U. N. Test			Test Instructions	
Formula	Raw Score	Interpretation	Raw Score	Interpretation
Dale-Chall	8.8625	Grades 11–12	6.7622	Grades 7–8
Flesch	45.3711	Grades 13–16	64.8980	Grades 8–9
Lorge	7.7462	Grades 7+	6.3335	Grades 6+

TABLE IV.—MEAN SCORES ON U. N. READING TEST

Subtests	No. of Items	Mean	Mean Percentage Correct
1	9	6.44	71.56
2	22	15.47	70.32
3	6	3.79	63.17
4	32	17.36	54.25
Total	69	43.06	62.41

able to perform on the U. N. reading test. Mean scores on each of the subtests are shown in Table IV.

In the absence of any absolute standard of how well students reading at the Grade XIII level should be able to perform on the U. N. reading test, it still appears reasonable to conclude that even at this level of ability the test offers considerable difficulty. It would seem that students who are reading below the Grade XIII level would experience even more trouble. Judged by this criterion of reading ability, the readability formulae appear to place the U. N. materials at too easy a level, some formulae showing a greater disparity between reading ability and readability than do others.

SUMMARY AND CONCLUSIONS

The present findings offer support to Lorge's contention that different formulae may be more suitable for some materials than for others. If this is so, there is need for research to deter-

mine the most appropriate formula to use with specific types of materials. There seems no doubt that factors other than those now being considered must be employed if suitable formulae are to be developed. As Lorge has pointed out, present techniques really do not assess the difficulty of the concepts found in the materials. Length of word and number of syllables are not necessarily correlated with abstractness of a concept, for some very abstract concepts involve few letters, whereas some concrete concepts are expressed by lengthy words.

The following conclusions appear to be justified by the evidence here presented:

1) There is considerable disparity between the readability levels found for U. N. materials by different readability formulae.

2) The Flesch and the Dale-Chall formulae estimate the readability of this type of material somewhat more accurately than does the Lorge formula, if level of reading ability is used as the criterion.

3) The U. N. materials sampled appear to be too difficult for high-school students.

Finally, it would appear that there is need for extended study of the readability of various types of materials intended for use in the public schools. These studies might well be carried out on regular types of school materials as well as on special materials such as pamphlets at various grade levels. Further study of the problem of readability is indicated, probably especially designed to consider the nature of the concepts involved. Lastly, it may be that the basic formulae are logically sound, and that the problem is the development of more valid norms.

BIBLIOGRAPHY

1) Julian Alden. "Lots of Names—Short Sentences—Simple Words," *Printers Ink*, 1945, 29.

2) Edgar Dale and J. S. Chall. "Formula for predicting readability: instructions," *Educational Research Bulletin*, 1948, 27, 11–20, 37–54.

3) J. N. Farr and J. J. Jenkins. "Tables for use with the Flesch Readability Formula," *Journal of Applied Psychology*, 1949, 33, 275–8.

4) Pauline J. Fihe, Viola Wallace and Martha Schulz, Compilers. *Books for Adult Beginners; Grades I to VII.* (Rev. Ed.) Chicago: American Library Association, 1946.

5) Rudolf Franz Flesch. "A dissenting opinion on readability," *Elementary English*, 1949, 26, 332–4.

6) ———. "A new readability yardstick," *Journal of Applied Psychology*, 1948, 32, 221–233.

7) ———. "Readability formula in practice," *Elementary English*, 1945, 25, 344–51.

8) ———. *The Art of Readable Writing.* New York: Harper and Brothers, 1949.

9) Alberta S. Gilinsky. "How Valid is the Flesch Readability Formula?" *American Psychologist*, 1948, 3, 261. (Abst.)

10) S. N. Gluckenheimer. "Readability of pamphlets on international relationships," *Educational Research Bulletin*, 1947, 26, 231–8.

11) Robert Gunning. "Gunning finds papers too hard to read," *Editor and Publisher*, 1945, May 19, 12.

12) P. M. Hayes, J. J. Jenkins and B. J. Walker. "Reliability of the Flesch Readability Formulas," *Journal of Applied Psychology*, 1950, 34, 22–26.

13) J. J. Jenkins. "Readability of written communications in industrial situations," *American Psychologist*, 1948, 3, 356. (Abst.)

14) Bryant Kearl. "A closer look at readability formulas," *Journalism Quarterly*, 1948, 25, 344–348.

15) Irving D. Lorge. "Lorge and Flesch Readability Formulae: Correction," *School and Society*, 1948, 67, 141–2.

16) ———. "Predicting readability," *Teachers College Record*, 1943–44, 45, 404–19.

17) ———. "Readability formulae: evaluation," *Elementary English*, 1949, 26, 86–95.

18) Melvin Lostutter. "Some critical factors of newspaper readability," *Journalism Quarterly*, 1947, 24, 307–314.

19) John U. Michaelis. "United Nations Reading Materials," *Social Education*, 1950, 14, 221–222.

20) Leo R. Miller. "Reading grade placement of the first twenty-three books awarded the John Newbery Prize," *Elementary School Journal*, 1946, 46, 394–399.

21) S. Pasholian and W. J. E. Crissy. "How readable are corporate annual reports?" *Journal of Applied Psychology*, 1950, 34, 244–248.

22) U. S. Department of Agriculture, Soil Conservation Service. *Our American Land; the story of its abuse and conservation.* (Misc. Publ. No. 596.) Washington, D. C.: U. S. Government Printing Office, 1946.

NOTES ON A SUGGESTED INDEX OF ITEM VALIDITY: THE U-L INDEX

A. PEMBERTON JOHNSON

Educational Testing Service
Princeton, New Jersey

In 1935 Long, Sandiford, et al,* reported an empirical comparative study of the effectiveness of thirteen validity techniques in selecting from among four hundred eighty-two six-choice vocabulary items that set of fifty and that set of one hundred items which correlated highest with criterion scores based on all four hundred eighty-two items. For this situation they concluded, among other things, that those validity techniques which tend to discriminate in favor of items of fifty per cent difficulty appear to be superior to the techniques such as bi-serial *r*, Kelley and Vincent which make no discrimination in favor of median difficulty items. Considering both effectiveness and ease of computation, they concluded that among those techniques tending to select items of fifty per cent difficulty (percentage passing), the upper vs. lower (group) technique was to be preferred. They stated that item validities derived by the upper vs. lower technique tend to reach maximum reliability when the proportion of cases in each of these groups is twenty-seven per cent.† Numerically the item validities were expressed in this technique by the difference between number of correct responses for the highest twenty-seven per cent (or by the highest third) and the number of correct responses for the lowest twenty-seven per cent (or the lowest third) of the cases as arrayed on the criterion scale.

This paper proposes the U-L Index, a simple modification of the upper-lower twenty-seven per cent technique. This index proved extremely useful during World War II as a basis for selecting items to be included in Army Air Forces bombardier knowledge tests. The U-L Index effectively tends to select

* John A. Long, Peter Sandiford, and Others. *The Validation of Test Items.* Bulletin No. 3, Dept. of Ed. Res., University of Toronto, Toronto, 1935, pp. 94–5.

† See also T. L. Kelley, "The Selection of Upper and Lower Groups for the Validation of Test Items," *J. Educ. Psychol.*, xxx, January, 1939, pp. 17–24.

items of moderate difficulty which show good separation between upper and lower twenty-seven per cent groups. It is more easily computed and interpreted than any conventional item validity index known to the author. For many practical test construction situations where expensive item analysis cannot be justified, it is believed to be quite useful.

This modification possibly has been used by other investigators before, but is believed not to have been. It requires the difference between the groups in number of correct responses to be divided by the number of cases in each group. Symbolically, if

R_u = number correct responses in the upper group of twenty-seven per cent

R_l = number correct responses in the lower group of twenty-seven per cent

N = number of cases

$f = .27N$

$$\text{U-L Index} = \frac{R_u - R_l}{.27N} = \frac{R_u - R_l}{f} \qquad (1)$$

This index can theoretically vary from +1 through 0 to −1. These proportions are independent of the N of the population sample and similar in magnitude though not strictly in meaning to conventional correlation coefficients.

Long, Sandiford, et al., found in their empirical study that (1) the average difficulty of a group of items shown to be least valid by the bi-serial *r* technique was about equal to the average difficulty of the group of items indicated to be least valid by the upper-lower thirds technique, but (2) that the bi-serial *r* tended to eliminate items indiscriminately throughout the entire range of item difficulty whereas the upper-lower thirds technique tended to eliminate items farthest removed from the fifty per cent level of difficulty (percentage passing).*

In a table of intercorrelations among the validity values obtained for one hundred items by the thirteen methods which they studied, Long, Sandiford, et al., report a correlation of .65 between bi-serial *r* and upper-lower thirds validities. This finding strongly suggests the need to interpret the proposed

* Ibid., p. 114.

U-L Index in somewhat different fashion than the conventional bi-serial r.

The standard error of the U-L Index may be derived from the conventional formula for the standard error of the difference of two proportions, when the data are uncorrelated.

Guilford* gives the following generalized formulae:

$$\#34 \qquad \sigma_{dp} = \sqrt{\sigma_{p_1}^2 + \sigma_{p_2}^2 - 2r\sigma_{p_1}\sigma_{p_2}}$$

$$\#28 \qquad \sigma_p = \sqrt{\frac{pq}{N}}$$

where: σ_{dp} = S.E. of a difference between two proportions
σ_p = S.E. of a proportion
p_1 = 1st proportion
p_2 = 2nd proportion
$q = 1 - p$
N = number of cases
r = coefficient of intercorrelation of the two sets of data.

Using the previously mentioned symbols for the U-L Index:

$$\sigma_{p_1} = \sqrt{\frac{p_1 q_1}{N}} = \sqrt{\frac{\left(\frac{R_u}{f}\right)\left(1 - \frac{R_u}{f}\right)}{f}} = \frac{1}{f}\sqrt{\frac{R_u(f - R_u)}{f}}$$

$$= \frac{1}{f}\sqrt{R_u - \frac{R_u^2}{f}}$$

Similarly $$\sigma_{p_2} = \frac{1}{f}\sqrt{R_l - \frac{R_l^2}{f}}$$

In Guilford's formula #34 for σ_{dp}, the most conservative estimate of r is 0. That formula then reduces to $\sigma_{dp} = \sqrt{\sigma_{p_1}^2 + \sigma_{p_2}^2}$

By substitution:

$$\text{SE U-L Index} = \frac{1}{f}\sqrt{R_u - \frac{R_u^2}{f} + R_l - \frac{R_l^2}{f}}$$

$$= \frac{1}{f}\sqrt{R_u + R_l - \left(\frac{R_u^2 + R_l^2}{f}\right)}$$

The formula for the standard error of the U-L Index is such that the maximum value of the standard error occurs in the

* J. P. Guilford. *Fundamental Statistics in Psychology and Education*, New York: McGraw-Hill, 1942, pp. 133 and 142.

TABLE I.—5% AND 1% "t" TEST SIGNIFICANCE LEVELS FOR THE SE OF U-L INDEX WHEN THE U-L INDEX = 0 AND 50% PASS THE ITEM*

N	$f = .27N$	SE of the U-L Index	5% Level	1% Level
100	27	.136	.273	.364
148	40	.112	.223	.295
200	54	.096	.190	.252
248	67	.087	.172	.227
300	81	.079	.156	.206
348	94	.073	.144	.190
400	108	.068	.134	.177
500	135	.061	.120	.159
1000	270	.043	.084	.111

* For example, only 1 per cent of obtained values would be expected to exceed ±.364 were the true index zero for a sample of one hundred cases. In this table the degrees of freedom have been taken as equal to $2(f\text{-}1)$ since f is defined as the number in each of the two contrasted groups.

fifty per cent passing situation where the true index equals zero. Table I below gives the 5% and 1% "t" significance levels of the SE of the U-L Index in that situation.

A quite serviceable index of item difficulty (percentage passing) for each item can be obtained from the values for R_u and R_l by use of the following formula:

$$\text{Difficulty level (percentage passing)} = \frac{100\ (R_u + R_l)}{2f} = \frac{50(R_u + R_l)}{f} \qquad (3)$$

Since this is based on only the extreme fifty-four per cent of the cases, it does not give exactly the value for the index of item difficulty which would be obtained from all of the cases. In practice however the discrepancy usually is not appreciable.

Formula (1) for the U-L Index provides no specific correction for omissions, although it is necessary sometimes to item analyze speeded tests in which the number of omissions among later items is appreciable.

Omissions on speeded tests are usually of two types: (a) those by persons who read the item but fail to mark any choice and

(b) those by persons who have not read as far in the test as the end of that item. Davis* in a communication to the writer suggests a method of determining the 'non-reads' directly. The test papers are scored and arranged in rank order according to the criterion score. The group which includes the top twenty-seven per cent of papers is segregated as is the group which comprises the bottom twenty-seven per cent of test papers. Next the papers of the upper group are gone through to determine which item on each paper is the last one marked. A tally is then made opposite the number of the next item for that one is presumably the first item not read by the subject. This tally is the basis for a cumulative frequency table of the number of cases within the upper group who have not read each particular item. A similar count is made of the 'non-reads' in the lower group or the bottom twenty-seven per cent of the test papers.

By graphic item count on the International Test Scoring machine, or by other means, the number of right responses and the number of wrong responses in the upper group to each item is found. The omits may be obtained, if desired, by subtracting from the number of papers within the upper group the rights plus wrongs plus 'non-reads.' Similarly the omits among the lower group may be obtained, if desired.

Where the 'non-reads' represent an appreciable proportion of either criterion group the following formula for the U-L Index is suggested:

$$\frac{R_u}{(f - \text{'non-reads'}_u)} - \frac{R_l}{(f - \text{'non-reads'}_l)} \qquad (4)$$

where R_u = number of right responses in the upper group
R_l = number of right responses in the lower group
$f = .27N$
'non-reads'$_u$ = 'non-reads' in the upper group
'non-reads'$_l$ = 'non-reads' in the lower group

The proportion $\frac{R_u}{(f - \text{'non-reads'}_u)}$ is the proportion of right responses in the upper group to the number in that group who

* Dr. Frederick B. Davis, formerly with the Coöperative Test Service and the Aviation Psychology Program of the Army Air Forces, now at Hunter College.

read the item. The proportion $\frac{R_t}{(f - \text{'non-reads'}_t)}$ may be similarly interpreted for the lower group.

The U-L Index is proposed as a useful means of selecting items either to increase the internal consistency (odd-even reliability) of a test or to increase its correlation with a given criterion. Some advantages of the U-L Index are: (1) the demonstrated effectiveness of the upper vs. lower twenty-seven per cent technique on which it is based, (2) its ease of computation, and (3) the readiness with which its level of significance may be approximated.

(*Concluded from Inside Front Cover*)

something like uniformity of style, footnotes and bibliographical references frequently require considerable 'marking for the printer,' and this is difficult—if not impossible—where matter is single-spaced. Manuscripts should not be marked for type style—this is done in the editorial office.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of the issue with which change is to take effect; both the old and new address should be given.

REGARDING CLAIMS FOR NON-RECEIPT

The Publisher desires every subscriber to get all issues to which he is entitled. Each journal is securely enclosed in a sturdy wrapper on which the subscriber's name and address have been printed, and is delivered directly to the Post Office, postage prepaid. Second-class matter is handled less by postal employees than other mail; moreover, if the Post Office is unable to make delivery, a notice to this effect is sent the Publisher and the magazine returned. Consequently, it is doubtful if one journal in a thousand is actually lost in transit.

But after an issue has been delivered to the proper address many things may happen to it—it may be diverted, or misplaced, or borrowed for 'keeps.' For this neither Post Office nor Publisher is responsible. However, too readily a subscriber who does not find a given issue in its assigned place makes a claim of non-receipt. Typical is the following from a Tennessee college librarian: "In checking through our periodical file we find that we did not receive January to May issues of *Educational Administration & Supervision* or March *Journal of Educational Psychology*. Please send these at once to complete our file." When informed that failure to locate an issue in the library was not evidence of non-receipt, he replied hotly that the Publisher was presumptuous in interpreting "periodical file" to mean library shelves instead of the "checking-in records."

In contrast—A periodical librarian in Iowa writes: "*The Journal of Educational Psychology* is too great temptation to our students. We cannot keep it on the shelves, so when it comes to time for binding little is left to bind. Please send with bill all 1949 and 1950 issues except those for December." From Minnesota: "Enclosed is $1.00 for replacement copy of December 1950 *Educational Administration & Supervision*. This issue disappeared very suddenly just as we are ready to send the volume to the bindery." A West Virginia college librarian writes: "We find, unhappily, that two issues of the last volume of the Journal are missing from our files. Apparently, our students thought they wanted them more than the library did. If you can supply January and October 1950 numbers, send with bill in duplicate."

Each subscriber is entitled to eight issues on a year's subscription. If he believes that of the eight issues mailed he got only five, it is his province to ascertain what became of the other three numbers. He should not be content with getting less than the eight issues to which he is entitled, but the Publisher cannot undertake to mail eleven—or even nine—issues in order that a subscriber may get his eight numbers.

A notice that claims for non-receipt of an issue should be made within two weeks after arrival of the next succeeding number has appeared regularly in each issue of the Journal. This rule will be adhered to rigidly in the future. In order that a claim may arrive within the time limit it should be addressed to the Publisher—not to an agency.

WARWICK AND YORK *Publishers* **BALTIMORE 2, MD.**

The
Journal
of
Educational
Psychology

INCHES
1
2
3
4
5
6
C2399
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
METRIC